



## A study on tensor and matrix models for super-resolution fluorescence microscopy

José Henrique de Morais Goulart, Laure Blanc-Féraud, Eric Debreuve,  
Sébastien Schaub

### ► To cite this version:

José Henrique de Morais Goulart, Laure Blanc-Féraud, Eric Debreuve, Sébastien Schaub. A study on tensor and matrix models for super-resolution fluorescence microscopy. CAMSAP 2019 - IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, Dec 2019, Le Gosier, Guadeloupe. hal-02309659

**HAL Id: hal-02309659**

**<https://hal.science/hal-02309659>**

Submitted on 9 Oct 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A STUDY ON TENSOR AND MATRIX MODELS FOR SUPER-RESOLUTION FLUORESCENCE MICROSCOPY

J. H. de M. Goulart, L. Blanc-Féraud, E. Debreuve

S. Schaub

Université Côte d’Azur  
CNRS, INRIA, I3S Laboratory, France  
{goulart, blancf, debreuve}@i3s.unice.fr

Université Côte d’Azur  
CNRS, Institut de Biologie Valrose, France  
schaub@unice.fr

## ABSTRACT

Super-resolution techniques for fluorescence microscopy are invaluable tools for studying phenomena that take place at sub-cellular scales, thanks to their capability of overcoming light diffraction. Yet, achieving sufficient temporal resolution for imaging live-cell processes remains a challenging problem. Exploiting the temporal fluctuations (blinking) of fluorophores is a promising approach that allows employing standard equipment and harmless excitation levels. In this work, we study a novel constrained tensor modeling approach that takes this temporal diversity into account to estimate the spatial distribution of fluorophores and their overall intensities. We compare this approach with an also novel matrix-based formulation which promotes structured sparsity via a continuous approximation of the cardinality function, as well as with other state-of-the-art methods.

**Index Terms**— fluorescence microscopy, super-resolution, tensor, structured sparsity.

## 1. INTRODUCTION

Because of light diffraction, the response of an optical microscope (point spread function, or PSF) forms a pattern known as Airy disk with a few hundred nanometers in diameter [1] which blurs nearby point sources together, limiting the spatial (lateral) resolution of conventional modern microscopes to around 200 nm. This limitation was first overcome in the 90’s and 2000’s with the advent of several super-resolution fluorescence microscopy techniques. Some of these techniques use specialized excitation laser beams that either inhibit emissions outside a small region of interest or display a periodic illumination pattern which reveals high-frequency contents of the sample [2,3]. Other approaches employ photoactivable or photoswitchable emitters (i.e., fluorescent molecules) and a careful manipulation of the excitation beam to acquire several images, each containing only a subset of active emitters [4].

In particular, single molecule localization microscopy (SMLM) techniques acquire many images with only a few

active molecules per image. The observed emissions are thus more likely to be isolated and can then be processed for an accurate localization of their centers, thereby allowing a fine reconstruction of the underlying sample [5]. SMLM techniques can reach resolutions down to 20 nm, but employ high-power excitation, which can be harmful for living cells [5]. Moreover, they require a precise operation of the excitation beam and a very time-consuming acquisition (taking from a few to many minutes). This severely limits the timescale of live-cell phenomena that can be investigated.

A less constraining approach consists in acquiring a temporal sequence of images over a significantly shorter time interval, so that each image displays a large subset of active emitters. Sparse deconvolution algorithms such as those of [6, 7] can then be leveraged to estimate emitters’ locations. Alternatively, one can exploit the *temporal diversity* induced by the blinking and bleaching (permanent depletion) of individual molecules. For instance, super-resolution optical fluctuation imaging (SOFI) computes (temporal) high-order cumulants of the sequence, effectively sharpening the PSF thanks to the statistical independence of emitters [8]. Super-resolution radial fluctuations (SRRF) microscopy also relies on computing statistics, but after performing a local (so-called “radiality”) transformation of each image in order to effectively reduce the PSF size [9]. However, these methods typically perform poorly for short image sequences.

In this paper we study a novel constrained tensor modeling approach that exploits spatial and temporal diversities and the knowledge of the PSF to estimate a map of local molecule densities of a 2D image and their overall intensities. In particular, we discuss and illustrate some limitations of this approach. Moreover, it is compared with an also novel matrix-based formulation which promotes structured sparsity via a continuous approximation of the  $\ell_0$  regularizer [10], as well as with other state-of-the-art methods [6,9].

## 2. SPATIOTEMPORAL IMAGE STACK MODEL

### 2.1. Molecule emission model

We consider a 2D ( $x$ - $y$  plane) sample model with  $R$  stationary fluorophores. Each emitter repeatedly and stochastically

This work was supported by the French Agence Nationale de la Recherche in the context of the project Investissements d’Avenir UCA<sup>JEDI</sup> (ANR-15-IDEX-01).

switches among emitting, dark and bleached states. Hence, by acquiring many image frames over an adequate timescale (typically up to a few seconds), the photon emission fluctuations exhibited by each molecule can be observed. Moreover, as molecules are small (measuring a few nanometers in  $x$  and  $y$ ) relative to the image size, the resulting time-varying emission can be modeled as a sum of weighted Dirac measures

$$I_0(x, y, k) = \sum_{r=1}^R s_r(k) \delta(x - x_r, y - y_r), \quad (1)$$

where  $(x_r, y_r)$  is the  $(x, y)$ -position of the  $r$ th molecule and  $s_r(k)$  is its intensity at the  $k$ th frame.

For simplicity, we discretize molecule locations over a high-resolution  $L \times L$  grid of  $\Delta \times \Delta$  pixels. Under this assumption, the  $k$ th frame can be expressed as the  $L \times L$  matrix

$$X_k := \sum_{r=1}^R s_r(k) e_{j_r} e_{i_r}^\top, \quad (2)$$

where  $(i_r, j_r) \in \{1, \dots, L\}^2$  is the discretized position of the  $r$ th molecule and  $\{e_l\}_{l=1}^L$  is the canonical basis of  $\mathbb{R}^L$ .

## 2.2. Matrix model of acquired stack

Assuming a linear time-invariant acquisition system is used, the  $k$ th frame results from the two-dimensional convolution of  $I_0(x, y, k)$  with the PSF  $h(x, y)$  that characterizes the system. For typical widefield microscopes, the 2D PSF in the radial plane can be well approximated by the Gaussian function

$$h(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right). \quad (3)$$

We will assume that  $\sigma_x = \sigma_y = \sigma$ .

Since the PSF satisfies  $h(x, y) = g(x)g(y)$ , where  $g$  is a zero-mean Gaussian of variance  $\sigma^2$ , we can express the  $k$ th acquired low-resolution  $N \times N$  frame in vectorized<sup>1</sup> form as

$$i_k = (G \boxtimes G) \text{vec}(X_k) \quad \left(\in \mathbb{R}^{N^2}\right) \quad (4)$$

$$= (G \boxtimes G) \sum_{r=1}^R s_r(k) (e_{i_r} \boxtimes e_{j_r}) \quad (5)$$

$$= (G \boxtimes G) \underbrace{(e_{i_1} \boxtimes e_{j_1} \dots e_{i_R} \boxtimes e_{j_R})}_{:=W} \begin{pmatrix} s_1(k) \\ \vdots \\ s_R(k) \end{pmatrix}, \quad (6)$$

where  $\boxtimes$  denotes the Kronecker product and  $G \in \mathbb{R}^{N \times L}$ , with  $N := L/q$  and  $q > 1$ , is a matrix whose  $l$ th column satisfies

$$(g)_n = \int_{(n-1)q\Delta}^{nq\Delta} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-(l-\frac{1}{2})\Delta)^2}{2\sigma^2}\right) dx \quad (7)$$

and represents the spatial PSF profile along  $x$  or  $y$  of an emitter centered at pixel  $l$ . From (7), (4) amounts to a convolution of (2) with a discretized Gaussian PSF followed by integration of every  $q$  consecutive pixels along each dimension.

<sup>1</sup>We define the vectorization of an  $N \times K$  matrix  $M$  as the stacking of its columns  $m_k$  into a long (column) vector  $\text{vec}(M) := (m_1^\top \dots m_K^\top)^\top$ .

By adjoining all vectorized frames described by (6) in a  $N^2 \times K$  matrix, we get

$$I = (i_1 \dots i_K) = (G \boxtimes G) W S^\top \in \mathbb{R}_+^{N^2 \times K}, \quad (8)$$

where  $(G \boxtimes G)$  is  $N^2 \times L^2$ ,  $W$  is  $L^2 \times R$  and  $S$  is a  $K \times R$  matrix with components  $(S)_{k,r} = s_r(k)$ .

## 2.3. Matrix-based formulation

Our problem boils down to estimating the nonnegative spatial and temporal profile matrices  $W$  and  $S$  from the acquired (noisy) data  $I$  and the knowledge of  $G$ . However, further assumptions are required, because the factorization is not unique. One can exploit the following priors:  $R$  is unknown but should satisfy  $R \ll L^2$ ;  $W$  and  $S$  are both nonnegative; and  $W$  has exactly one nonzero element per column. Relaxing the latter prior into a sparsity assumption leads to a regularized least-squares formulation of the form

$$\min_{W \geq 0, S \geq 0} \frac{1}{2} \|I - (G \boxtimes G) W S^\top\|_F^2 + \alpha \|W\|_1 + \gamma \|S\|_{2,1}, \quad (9)$$

where  $\|S\|_{2,1} := \sum_{r=1}^R \|s_r\|_2$ , with  $s_r$  denoting the  $r$ th column of  $S$ , and  $\|W\|_1 := \sum_{l=1}^{L^2} \sum_{r=1}^R |w_{l,r}|$ . The term  $\|W\|_1$  encourages sparsity in  $W$ , as in the lasso problem [11]. Owing to its geometric properties, the  $\ell_{2,1}$ -norm  $\|S\|_{2,1}$  promotes column-wise sparsity of  $S$ : as  $\gamma$  grows, whole columns of  $S$  are nullified, reducing the effective number of molecules.

A major drawback of the above formulation is the number of variables, which grows as  $\mathcal{O}((L^2 + K)R)$ . This issue is alleviated by using a tensor approach, as discussed next.

## 3. REGULARIZED TENSOR DECOMPOSITION APPROACH

### 3.1. Tensor-based formulation

A natural way to reduce the number of unknowns in (8) is by exploiting the tensor structure of the columns of  $W$ . Writing  $w_r = v_r \boxtimes u_r$  with  $u_r, v_r \in \mathbb{R}_+^L$  leads to the tensor model

$$\mathcal{J} = \sum_{r=1}^R G u_r \otimes G v_r \otimes s_r \in \mathbb{R}_+^{N \times N \times K}, \quad (10)$$

where  $s_r$  is the  $r$ th column of  $S$  and  $\otimes$  denotes the outer (tensor) product [12]. Concretely, the  $r$ th summand of (10) is an  $N \times N \times K$  tensor with components  $(G u_r \otimes G v_r \otimes s_r)_{ijk} = (G u_r)_i (G v_r)_j (s_r)_k$ . Expression (10) is known as polyadic (or PARAFAC) decomposition (PD) of tensor  $\mathcal{J}$ , and is multilinear with respect to its matrix factors  $U := (u_1 \dots u_R)$ ,  $V := (v_1 \dots v_R)$ , both of size  $L \times R$ , and  $S$ . A recent account on this decomposition and its properties is given in [12].

Analogously to (9), model (10) leads to the formulation

$$\min_{U, V, S \geq 0} f(U, V, S) + \alpha (\|U\|_1 + \|V\|_1) + \gamma \|S\|_{2,1} \quad (11)$$

where  $f(U, V, S) \triangleq \frac{1}{2} \|\mathcal{J} - \sum_{r=1}^R G u_r \otimes G v_r \otimes s_r\|^2$ . The number of variables of this formulation grows as  $\mathcal{O}((2L +$

$K)R$ ). As in (9), the parameter  $\gamma$  allows trading data fidelity for parsimony with respect to the number of molecules. We note that  $\ell_{2,1}$ -norm penalization has already been used in [13, 14] for promoting parsimony of tensor models.

### 3.2. Optimization algorithm

Problem (11) is nonconvex and nonsmooth, and thus considerably difficult. We develop next a variant of the alternating group lasso (AGL) algorithm [14], which is based on the observation that minimizing (11) with respect to only one of the three variables  $U$ ,  $V$  and  $Z$  is a convex (though still nonsmooth) problem. Specifically, by fixing  $V = \hat{V}$  and  $S = \hat{S}$  and vectorizing the tensor model, one obtains the subproblem

$$\min_{u \in \mathbb{R}_+^{LR}} \frac{1}{2} \|i^{(u)} - [(\hat{S} \odot G\hat{V}) \boxtimes G] u\|^2 + \alpha \|u\|_1, \quad (12)$$

where  $\odot$  is the Khatri-Rao (columnwise Kronecker) product,  $i^{(u)}$  is the vectorization of tensor  $\mathcal{J}$  with elements  $i_{i+N(j-1)+N^2(k-1)}^{(u)} = (\mathcal{J})_{i,j,k}$  and  $u = \text{vec}(U)$ . Subproblem (12) is a standard lasso problem [11] and thus can be solved by many existing algorithms. A formally identical subproblem in  $V$  can be derived by fixing  $U = \hat{U}$  and  $S = \hat{S}$ , but this time the data vector  $i^{(v)}$  has elements  $i_{j+N(i-1)+N^2(k-1)}^{(v)} = (\mathcal{J})_{i,j,k}$  and the matrix multiplying  $v$  is  $(\hat{S} \odot G\hat{U}) \boxtimes G$ . Similarly, the subproblem in  $S$  is a group lasso problem [11] with disjoint groups, and can be written as

$$\min_{s \in \mathbb{R}_+^{KR}} \frac{1}{2} \|i^{(s)} - [(G\hat{V} \odot G\hat{U}) \boxtimes E_K] s\|^2 + \gamma \sum_{r=1}^R \|s_r\|_2,$$

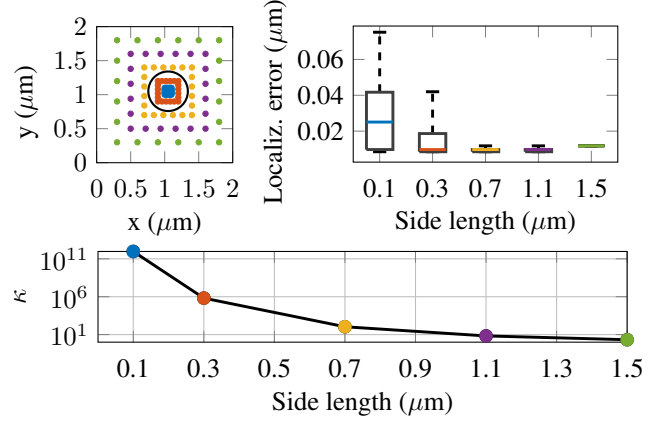
where  $E_K$  is the  $K \times K$  identity matrix.

Hence, at each iteration one sequentially solves the three described subproblems with the fixed variables set to their current estimates. To guarantee convergence, one can add a proximal term to the cost function of each subproblem. Then, by using the same arguments as in [14], it can be shown that this algorithm converges to a stationary point of (11).

### 3.3. Discussion

Without constraints, the tensor model in (10) is not identifiable: two molecules can have e.g. an identical  $x$  coordinate, and thus  $G u_r = G u_p$  for some distinct indices  $r$  and  $p$ , implying  $G u_r \otimes G v_r \otimes s_r + G u_p \otimes G v_p \otimes s_p = G u_r \otimes (G v_r \otimes s_r + G v_p \otimes s_p)$ . As the matrix between parentheses in this expression equals  $G(v_r \ v_p) T T^{-1} (s_r \ s_p)^T$  for any  $2 \times 2$  nonsingular matrix  $T$ , it follows that the corresponding spatial and temporal profiles cannot be identified.

Identifiability can be potentially restored by enforcing non-negativity and sparsity of  $u_r$ ,  $v_r$  and incorporating knowledge of  $G$ . To our knowledge, there are no theoretical results confirming this observation (the closest ones seem to be those in [15], where a PD with a dictionary is considered). Nonetheless, our experiments indicate that molecule locations can indeed be recovered using the described algorithm.



**Fig. 1.** Results of tensor approach for square-shaped distributions. Top left: molecule positions; top right: localization error as a function of the square side length; bottom: condition number  $\kappa$  of matrix  $(G\hat{V} \odot G\hat{U}) \boxtimes E_K$  at first iteration. The colors used in all plots refer to their respective side lengths. The circle in the top left plot shows the FWHM of the PSF.

However, its performance degrades as the super-resolution factor  $q$  grows and the molecules get closer because then the matrix  $(G\hat{V} \odot G\hat{U}) \boxtimes E_K$  appearing in the subproblem in  $S$  becomes severely ill-conditioned.

This behavior is illustrated by the results of Fig. 1, where five stacks of 20 molecules uniformly distributed over the border of a square with varying side length are processed. The temporal profiles were generated with the SOFI simulation tool [16] and are discretized realizations of a continuous-time Markovian process with emitting (average lifetime  $\tau_{\text{on}} = 20$  ms), dark ( $\tau_{\text{off}} = 40$  ms) and bleached (average emitting time until bleaching  $\tau_{\text{bl}} = 80$  s) states. The number of photons per frame of an emitting molecule is a Poisson variable with mean 400. The PSF has a full width at half maximum (FWHM) of about 261 nm. We set  $q = 6$ , initialize the molecule positions with their true values and start each iteration by solving for  $S$ . The FISTA (fast iterative shrinkage-thresholding algorithm) scheme of [17] was used to solve the subproblems. As the molecules approach each other, localization precision is compromised. This is explained by the condition number  $\kappa$  (see Fig. 1) of  $(G\hat{V} \odot G\hat{U}) \boxtimes E_K$  at the first iteration.

## 4. ROW-SPARSE MATRIX APPROACH

### 4.1. Convex and nonconvex formulations

Another approach for recovering the super-resolved image is to estimate the product  $X := W S^T \in \mathbb{R}_+^{L^2 \times K}$  from the data while requiring that its number of nonzero rows be small, which leads to the following convex group lasso problem:

$$\min_X \frac{1}{2} \|I - (G \boxtimes G) X\|^2 + \alpha \|X\|_{1,2}, \quad (13)$$

with  $\|X\|_{1,2} := \sum_{l=1}^{L^2} \|e_l^\top X\|_2$ . The price to pay for this “convexification” is the larger number of variables, which now grows as  $\mathcal{O}(L^2 K)$ . On the other hand, the conditioning of  $G \boxtimes G$  depends only on  $(\sigma, q)$ , and not on molecule locations.

Alternatively, one can use a nonconvex regularizer to overcome the typical bias of solutions of (13); namely, the overall intensities of molecules (which correspond to the  $\ell_1$ -norms of rows of  $X$ ) tend to be underestimated. A well-suited option is to apply the continuous exact  $\ell_0$  (CEL0) functional [10] defined as  $\Phi(\bar{x}, \alpha) := \sum_{i=1}^{L^2} \phi(\bar{x}_i, \alpha)$ , where

$$\phi(\bar{x}_i, \alpha) := \begin{cases} \alpha, & |\bar{x}_i| \leq \frac{\sqrt{2\alpha}}{\|\bar{g}_i\|_2}, \\ \alpha - \frac{\|\bar{g}_i\|_2^2}{2} \left( |\bar{x}_i| - \frac{\sqrt{2\alpha}}{\|\bar{g}_i\|_2} \right)^2, & \text{otherwise,} \end{cases}$$

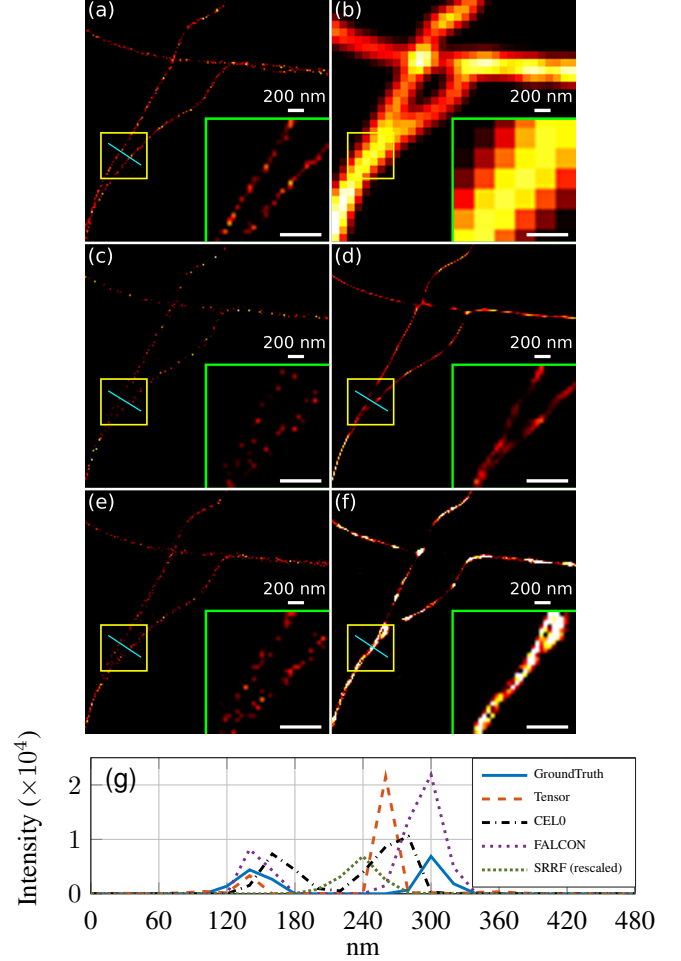
to the vector  $\bar{x} := (\|e_1^\top X\|_2 \dots \|e_{L^2}^\top X\|_2)^\top \in \mathbb{R}_+^{L^2}$ . CEL0 has the benefits of being continuous (unlike the cardinality function<sup>2</sup>  $\|\bar{x}\|_0$ ) and of preserving the global minima of the (much harder) problem  $\min_X \frac{1}{2} \|I - (G \boxtimes G) X\|^2 + \alpha \|\bar{x}\|_0$ , while eliminating some of its local minima [10].

## 4.2. Relation to existing approaches

Existing sparse deconvolution approaches such as FALCON [6] and SPIDER [7] estimate molecule locations for each frame separately. FALCON uses a  $\ell_1$  penalty and refines the estimated positions with a gradient descent algorithm; SPIDER uses  $\|x_k\|_0$ . To our knowledge, no work has yet employed a structured sparsity formulation to estimate the whole super-resolved stack at once as described above.

## 5. NUMERICAL RESULTS AND DISCUSSION

We simulated the acquisition of  $K = 10$  images of size  $N = 64$  over 0.1 sec using the molecule distribution of the microtubules dataset (MT0) from the SMLM challenge 2016,<sup>3</sup> containing 8 731 molecules. The temporal blinking profiles were generated using the SOFI simulation tool [16], as described in Section 3.3. The ground truth and observed image (with a  $5\times$  zoom) of Fig. 2(a)–(b) were computed by summing the corresponding frames. The PSF FWHM is approximately 261 nm. Fig. 2(c)–(d) show the sum of frames produced by the tensor and CEL0 approaches with  $q = 5$  (the fine-grid pixel size is  $\Delta = 20$  nm). The former is initialized by randomly placing at every (coarse-grid) pixel a number of molecules proportional to its intensity, and the latter by first solving (13) with a FISTA algorithm. The solution of CEL0 is computed with an iterative reweighted  $\ell_1$  algorithm [18] using FISTA at every step. The outcomes of FALCON and SRRF<sup>4</sup> (using second-order statistics of the radially maps) are shown in Fig. 2(e)–(f). Finally, Fig. 2(g) depicts the intensity profiles obtained for each



**Fig. 2.** Results for simulated microtubules dataset: (a) ground truth; (b) integrated observed stack ( $5\times$  zoom); (c) tensor approach; (d) CEL0; (e) FALCON; (f) SRRF; (g) intensity profiles along the shown blue line. The frame in the bottom right corner shows a  $2.66\times$  zoom of the smaller yellow frame.

method along the line inside the small square; the curve for SRRF has been rescaled to match the closest peak.

The tensor approach recovers the target structure but cannot reconstruct close fluorophores, while CEL0 yields the sharpest image. Computing times are 62.8, 87.9, 7.4 and  $<1$  sec for respectively the tensor, CEL0, FALCON and SRRF methods. The Jaccard index, which measures the accuracy of molecule detection and is defined as the ratio of true positives to the sum of true positives, false positives and false negatives, is respectively 0.27, 0.47, 0.33 and 0.35 for these methods.<sup>5</sup> In conclusion, despite the difficulty of this example where only 10 frames are available, the tensor, CEL0 and FALCON methods are able to reveal the overall target structure.

<sup>2</sup>Recall that  $\|x\|_0$  is defined as the number of nonzero components of  $x$ .

<sup>3</sup><http://bigwww.epfl.ch/smlm/datasets/index.html>.

<sup>4</sup>We have used the NanoJ SRRF plugin for ImageJ, available at: <https://github.com/HenriquesLab/NanoJ-SRRF>.

<sup>5</sup>We used the the Gale–Shapley algorithm [19] to match estimated and true molecules; the preference order is determined by distance, within a 20 nm tolerance radius. In each estimated image, pixels with intensity below 1% the image peak were discarded before the matching.

## 6. REFERENCES

- [1] G. Chartier, *Introduction to optics*, Springer Science & Business Media, New York, 2005.
- [2] S. W. Hell and J. Wichmann, “Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy,” *Optics Letters*, vol. 19, no. 11, pp. 780–782, 1994.
- [3] M. G. L. Gustafsson, “Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy,” *Journal of Microscopy*, vol. 198, no. 2, pp. 82–87, 2000.
- [4] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, “Imaging intracellular fluorescent proteins at nanometer resolution,” *Science*, vol. 313, no. 5793, pp. 1642–1645, 2006.
- [5] J. R. Allen, S. T. Ross, and M. W. Davidson, “Single molecule localization microscopy for superresolution,” *Journal of Optics*, vol. 15, no. 9, pp. 094001, 2013.
- [6] J. Min, C. Vonesch, H. Kirshner, L. Carlini, N. Olivier, S. Holden, S. Manley, J. C. Ye, and M. Unser, “FALCON: fast and unbiased reconstruction of high-density super-resolution microscopy data,” *Scientific Reports*, vol. 4, pp. 4577, 2014.
- [7] S. Hugelier, J. J. De Rooi, R. Bernex, S. Duwé, O. Devos, M. Sliwa, P. Dedecker, P. H. C. Eilers, and C. Ruckebusch, “Sparse deconvolution of high-density super-resolution images,” *Scientific Reports*, vol. 6, pp. 21413, 2016.
- [8] T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, “Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI),” *Proceedings of the National Academy of Sciences*, vol. 106, no. 52, pp. 22287–22292, 2009.
- [9] N. Gustafsson, S. Culley, G. Ashdown, D. M. Owen, P. M. Pereira, and R. Henriques, “Fast live-cell conventional fluorophore nanoscopy with ImageJ through super-resolution radial fluctuations,” *Nature Communications*, vol. 7, pp. 12471, 2016.
- [10] E. Soubies, L. Blanc-Fraud, and G. Aubert, “A continuous exact  $\ell_0$  penalty (CEL0) for least squares regularized problem,” *SIAM Journal on Imaging Sciences*, vol. 8, no. 3, pp. 1607–1639, 2015.
- [11] T. Hastie, R. Tibshirani, and M. Wainwright, *Statistical learning with sparsity: the lasso and generalizations*, Chapman and Hall/CRC, New York, 2015.
- [12] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, “Tensor decomposition for signal processing and machine learning,” *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3551–3582, 2017.
- [13] X. Han, L. Albera, A. Kachenoura, H. Shu, and L. Senghadji, “Block term decomposition with rank estimation using group sparsity,” in *Proceedings of the 2017 IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Curaçao, Dutch Antilles, Dec. 2017.
- [14] J. H. de M. Goulart, P. M. R. Oliveira, R. C. Farias, V. Zarzoso, and P. Comon, “Alternating group lasso for block-term tensor decomposition with application to ECG source separation,” *HaL preprint hal-01899469*, 2018, (in review).
- [15] J. E. Cohen and N. Gillis, “Dictionary-based tensor canonical polyadic decomposition,” *IEEE Transactions on Signal Processing*, vol. 66, no. 7, pp. 1876–1889, 2017.
- [16] A. Girsault, T. Lukes, A. Sharipov, S. Geissbuehler, M. Leutenegger, W. Vandenberg, P. Dedecker, J. Hofkens, and T. Lasser, “SOFI simulation tool: a software package for simulating and testing super-resolution optical fluctuation imaging,” *PLoS ONE*, vol. 11, no. 9, pp. e0161602, 2016.
- [17] B. O’Donoghue and E. Candès, “Adaptive restart for accelerated gradient schemes,” *Foundations of Computational Mathematics*, vol. 15, no. 3, pp. 715–732, 2015.
- [18] P. Ochs, A. Dosovitskiy, T. Brox, and T. Pock, “On iteratively reweighted algorithms for nonsmooth nonconvex optimization in computer vision,” *SIAM Journal on Imaging Sciences*, vol. 8, no. 1, pp. 331–372, 2015.
- [19] David Gale and Lloyd S Shapley, “College admissions and the stability of marriage,” *The American Mathematical Monthly*, vol. 69, no. 1, pp. 9–15, 1962.